12-2017

Reliability of an Observational Method Used to Assess Tennis Serve Mechanics in a Group of Novice Raters

Natalie L. Myers
University of Kentucky, nmyers02@gmail.com

W. Ben Kibler
Shoulder Center of Kentucky

Gilson J. Capilouto
University of Kentucky, gilson.capilouto@uky.edu

Philip M. Westgate
University of Kentucky, philip.westgate@uky.edu

Tony English
University of Kentucky, tenglish@uky.edu

See next page for additional authors

Reliability of an Observational Method Used to Assess Tennis Serve Mechanics in a Group of Novice Raters

Natalie L. Myers PhD, ATC, PES
W. Ben Kibler MD, FACSM
Gilson J. Capilouto PhD, FASHA
Philip Westgate, PhD
Tony English, PhD, PT
Tim L. Uhl PhD, ATC, PT, FNATA

ABSTRACT

Background: *Previous research has developed an observational tennis serve analysis (OTSA) tool to assess serve mechanics. The OTSA has displayed substantial agreement between the two health care professionals that developed the tool; however, it is currently unknown if the OTSA is reliable when administered by novice users.*

Purpose: *The purpose of this investigation was to determine if reliability for the OTSA could be established in novice users via an interactive classroom training session.*

Methods: *Eight observers underwent a classroom instructional training protocol highlighting the OTSA. Following training, observers participated in two different rating sessions approximately a week apart. Each observer independently viewed 16 non-professional tennis players performing a first serve. All observers were asked to rate the tennis serve using the OTSA. Both intra and inter-observer reliability were determined using Kappa coefficients.*

Results: *Kappa coefficients for intra and inter-observer agreement ranged from 0.09 to 0.83 depending on the body position. A majority of all body positions yeilded moderate agreement and higher.*

Conclusion: *This study suggests that the majority of components associated with the OTSA are reliable and can be taught to novice users via a classroom training session.*

Key Words: *tennis serve, observational analysis, reliability analysis*

INTRODUCTION

The tennis serve is the most predominant stroke during the service game and is thought to be the most important shot as it initiates the start of each point.[1] The serve is used as a weapon to dictate the point between two opponents. The execution of a perfect serve involves dynamic function of the entire kinetic chain. It requires a sequence of coordinated movements that requires the transfer of energy from the lower limbs to the upper limbs in a period lasting approximately 1 second.[2] As such, serve speeds in an elite population may reach up to 100 miles per hour with rotational velocities and torques about the Glenohumeral joint reaching up to 2420°/sec and approximately 55Nm/BW*H respectively, depending on the phase of motion.[2-4] Biomechanical alterations throughout the kinetic chain have been discovered between players with and without shoulder pathology using three dimensional (3D) motion analysis. Martin et al,[5] followed male tennis players for two seasons, and found that players who went on to sustain a shoulder injury demonstrated larger upper extremity joint kinetics, and decreased ball velocities compared to those who were injury free. This study indicates that the evaluation of tennis serve mechanics is important to potentially reduce injury risk.[5] However, access to 3D equipment is not always feasible for all coaches and health care professionals (HCP). Consequently, a field based method was developed to assess joint position and motion during the serve[6] that may eventually allow coaches and HCPs to evaluate mechanics for potential injury risk and performance flaws in the absence of expensive equipment.

Researchers investigating the biomechanical demands of the tennis serve most often utilize traditional three-dimensional motion analysis capture. The use of 3D motion capture is the gold standard in movement analysis;[7] however, the technology has limitations. Motion capture is costly, time consuming, and cannot be easily transported or utilized on court.[8-10] Thus, a field based method would offer an alternative for coaches and HCPs to evaluate tennis serve mechanics with the goal of improving performance and potentially diminishing injury risk. Such a tool would provide a cost-effective means of offering feedback to players on good and poor positioning during the service motion, so they could consider strategies to improve poor positioning. To that end, the Women's Tennis Association (WTA) and the Shoulder Center of Kentucky developed the observational tennis serve analysis (OTSA) tool for on court assessment of serve mechanics without the need for expensive laboratory equipment.[6]

The OTSA assesses key body positions and motions throughout the kinetic chain. The inter-observer reliability of the OTSA has been previously tested.[6] Reliability of the OTSA ranged from Kappas (K) between 0.36-1.0, with the majority of the body positions displaying substantial reliability ($K > 0.61$). These Kappa results suggest an acceptable agreement between the two health care professionals that created the OTSA. While these results are valuable, the tools external validity is lacking; more specifically, it is difficult to determine if this method could be used by coaches and health care professionals unfamiliar with the observational method.

Previous research investigating the reliability of observational methods of movement patterns have determined that an educational training session is imperative to yield superior reliability between novice raters. Self-instruction slide presentations and instructional compact discs have yielded moderate to excellent reliability when used in studies detecting scapular dyskinesis and knee valgus motion, respectively.[11,12] While self-instruction has been generally successful, other authors have suggested that incorporating more intensive training programs such as interactive classroom design, might also be useful.[13]

Therefore, the purpose of this study was to determine if reliability for the OTSA could be established in novice users via an interactive classroom training session. It was hypothesized that the reliability for all novice users (tennis coaches and HCPs) would be moderate ($Kappa \geq 0.41$) or higher for the majority of body positions associated with the OTSA.

METHODS

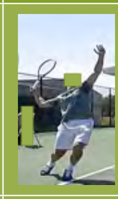
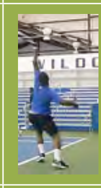
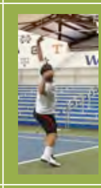
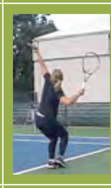
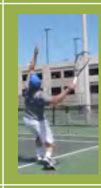
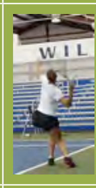
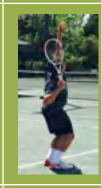
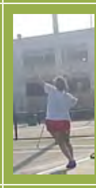
The University of Kentucky Institutional Review Board approved this study. Eight observers were recruited from a sample of convenience and underwent an OTSA classroom instructional training protocol along with two OTSA rating sessions. The observers viewed videos of 33 non-professional tennis players performing a first serve. The observers included 4 tennis coaches and 4 HCPS. Of the 4 coaches, 2 were high school (64 years of combined experience) and 2 were tennis professionals (21 years of combined experience). Of the HCPs, 2 were athletic trainers (combined 17 years of experience) and 2 were physical therapists (combined 16 years of experience). Tennis coaches were included if they were actively coaching at the recreational, high school, or college level. Retired coaches were able to participate if they had tennis coaching experience lasting longer than 10 years. HCPs were included it they were a certified athletic trainers (ATC) or licensed physical therapist (PT). Each player in the video was verbally informed of the study and voluntarily signed an informed consent form if over the age of 18 or assent form if under the age of 18.

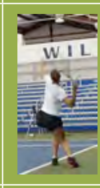
OBSERVATIONAL TENNIS SERVE ANALYSIS (OTSA) TOOL

Players' serves were captured using two digital cameras. The first camera was positioned anteriorly to the participant, 20 feet from the baseline "T" of the court at a 20° angle. The second was positioned posterolaterally to the participant, 14 feet from the baseline "T" of the court at a 45° angle.[6]

The OTSA has nine components. The first eight components are evaluated at maximal knee flexion while the last motion is assessed during the entire service motion. The first eight components are defined as nodes, and represent a body position at a specific joint. The nodes were compiled through 3D motion analysis studies.[4,5,14-20] The OTSA is accompanied by operational definitions that describe what is considered "good' and "bad" mechanics for each of the nine components (*Table 1*). For this study, the operational definition for node two and seven were altered in hopes of eliciting improved reliability from the original study.

Table 1. Observational Tennis Serve Analysis Tool Grading Scale

	Efficient Mechanics	Picture of Good Mechanics	Inefficient Mechanics	Picture of Bad Mechanics
Node 1: Foot	Good: Back foot stays behind front foot		Bad: Back foot stays in front of front foot	
				
Node 2: Knee	Good: Substantial knee bend (both knees bend >15°		Bad: None to minimal knee bend (both knees bend less than or equal to 15°	
Node 3: Counterhip Rotation	Good: The hip on back side is rotating away from the net		Bad: The hip on back side is not rotating away from the net	
Node 4: Posterior hip tilt	Good: The hip on back side is dropping towards the ground		Bad: The hip on back side is not dropping towards the ground	
Node 5: Hip Lean	Good: The hip on front side is not leaning forward towards the net		Bad: The hip on front side is leaning forward towards the net	
Node 6: X-Angle	Good: The shoulder rotate past the hips (x-angle = 30°)		Shoulders don't rotate behind the hips Bad: the x-angle is less than 30°	
			Shoulders rotate too far behind the hips Bad: the x-angle is greater than 30°	

Node 7: Trunk	Good: Trunk rotation around a vertical axis		Bad: No trunk rotation around a vertical axis	
Node 8: Arm	Good: Shoulder in line with the plane of scapula		Bad: Hypercocking – shoulder behind plane of scapula	
			Hypococking – shoulder in front of plane of scapula	
Assessment of Motion 9: Composite Motion of Kinetic Chain	Good: Used knee flexion and back leg drive to maximize ground reaction forces that push the body upward from the cocking position into ball impact	 Picture represents end stage of motion (motion to be assessed dynamically)	Bad: Use trunk muscles to pull the trunk and arm from cocking into ball impact	 Picture represents end stage of motion (motion to be assessed dynamically)
*Note: Evaluate nodes 1–8 at maximum knee bend. Composite motion of kinetic chain should be evaluated throughout entire motion.				

PROCEDURES

The lead author (NM) led the classroom instructional training session. Scheduling conflicts prohibited three coaches from attending the initial session, they were given the identical training session on a different day. The training session took place in a typical conference style room and included an hour and fifteen-minute interactive PowerPoint presentation of the OTSA tool followed by an initial rating session. The training session included background information on the development of the OTSA, information regarding the significance of the analysis method, detailed rating instructions, practice rating session for each individual node using picture and video examples from 17 player videos, and a final video assessment using the OTSA to grade tennis serve mechanics. To participate in the rating session, coaches and HCPs had to receive a 75% or better on the final assessment; all observers met the criteria. Observers were encouraged to ask questions and were permitted to share their decisions during the practice session. Any disagreement between the observers was discussed until a consensus could be reached.

The first rating session lasted approximately one hour and commenced once all observed felt confident with the instructions. Sixteen individual player videos were projected, different from those used in the instructional training session, onto a screen in the same conference room as the training session. Each video captured 1 service trial, and observers were asked to grade the trial using a standardized score sheet. The score sheet allowed observers to categorize each node in a binomial format as either “good” or “bad.” Each player video was identically edited to result in three different parts in the following order: 1) the observers were presented with an anterior view of the service motion and promoted to evaluate node 1 (5 second freeze frame at maximum knee bend); 2) from the posterior view observers had one minute to evaluate nodes 2-8 (video was freeze framed at maximum knee bend); and, 3) from the posterior view observers had 10 seconds to evaluate the composite motion of the kinetic chain (motion 9) while viewing a slow motion video of the entire serve.

Observers were given the following instructions prior to the start of the rating session: 1) observers could request multiple viewings, and 2) observers could not share their ratings or make any comments. To evaluate intra-observer reliability, observers were provided with a Universal Serial Bus (USB) drive to view the videos on their own computers. Observers assessed the same video footage one week later (range 7-22 days) to reduce the likelihood observers would remember their initial scores. The videos in the second viewing were presented in a different order from the first rating session. Each observer had access to a printed document that identified the operational definitions and picture representations describing both "good" and "bad" mechanics for all nine components associated with the OTSA during the training session, and for the initial and follow-up rating sessions.

STATISTICAL ANALYSIS

Unweighted Kappa (K) coefficients were used to determine intra-observer between day reliability for each of the nine components.[21] Fleiss's multi-rater Kappa coefficient was used to investigate inter-observer agreement on day 1 for all components of the OTSA. This statistic has been recommended for measuring agreement amount for two or more raters.[22,23] The multi-rater K from day 2 was not reported; kappas between day 1 and day 2 for each node were similar as determined by a two-sample Wald test ($p>0.05$). All Fleiss multi-rater Kappas were generated using an online calculator.[24] For the purposes of this study, the following scale was used to determine the strength of agreement between two raters: ≤ 0 = poor agreement, .01-.20 = slight agreement, .21-.40 = fair agreement, .41-.60 = moderate agreement, .61-.80 = substantial agreement, and .81-1 = almost perfect agreement.[25]

RESULTS

Intra-observer reliability

The average kappa values among the 8 observers was moderate and higher for all 9 components of the OTSA (*Table 2*).

Table 2: Intra-observer reliability for each of the nine components of the OTSA

Nodes	All 8 Observers	Kappa Interpretation
1	0.83 ± 0.10	Almost perfect agreement
2	0.78 ± 0.14	Substantial agreement
3	0.73 ± 0.15	Substantial agreement
4	0.61 ± 0.22	Substantial agreement
5	0.42 ± 0.26	Moderate agreement
6	0.57 ± 0.22	Moderate agreement
7	0.63 ± 0.21	Substantial agreement
8	0.65 ± 0.20	Substantial agreement
Motion 9	0.66 ± 0.18	Substantial agreement

Data represents the averaged kappa ± standard error for all 8 observers for each node.

Inter-observer reliability

Multi-rater kappa values were moderate agreement or higher for 8 out of the 9 components between all 8 observers. Node 5 generated slight agreement among the novice users (*Table 3*).

Table 3: Inter-observer reliability for each of the nine components of the OTSA

Nodes	All 8 Observers	Kappa Interpretation
1	0.64 ± 0.047	Substantial agreement
2	0.72 ± 0.047	Substantial agreement
3	0.57 ± 0.047	Moderate agreement
4	0.47 ± 0.047	Moderate agreement
5	0.09 ± 0.047	Slight agreement
6	0.49 ± 0.047	Moderate agreement
7	0.62 ± 0.047	Substantial agreement
8	0.43 ± 0.047	Moderate agreement
Motion 9	0.46 ± 0.047	Moderate agreement

Data represents multi-rater kappa ± standard error for all 8 observers for each node.

DISCUSSION

The purpose of this study was to determine if reliability for administering the OTSA could be established in a group of novice users via interactive classroom training. It was hypothesized that the reliability for all novice users (tennis coaches and HCPs) would be moderate (Kappa ≥ 0.41) or higher for the majority of body positions associated with the OTSA. Results support our research hypothesis; we found reliable assessment for the majority of nodes among novice users of the tool. Intra-observer reliability yielded higher kappa coefficients as compared to inter-observer reliability. Forward hip lean (node 5) was the weakest node for inter-observer reliability assessment, with values only reaching slight agreement.

Forward hip lean seems to be the most difficult node for both expert[6] and novice users to reliably judge. The initial study investigating the reliability of the OTSA among expert observers showed that forward hip lean generated the lowest kappa value (0.36) between the developers of the OTSA.[6] Despite the low agreement, this node remained a component of the OTSA so the authors could determine if incorporating a classroom training session could possibly improve reliability. The fact that our training resulted in equally low agreement suggests forward hip lean (node 5) may need to be removed from the OTSA. Whether expert or novice, observers are not able to consistently agree on whether or not the front hip leans forward during maximal knee bend. A possible explanation might have been due to camera positioning. Observers are asked to grade this particular node from a posterolateral vantage point which may not be optimal for visual assessment. Experimenting with other camera angles could improve reliability between multiple raters. Alternatively, better operationalization of what constitutes forward hip lean might be necessary to improve observer reliability.

Observational analysis of movement dates back to the early 1970s. Investigators found that experienced physical therapists agree on sagittal trunk and knee motion deviations during gait in adult hemiplegic patients with 93% accuracy.[26] Visual analysis is the most common approach to providing an estimation of kinematics,[8,27] and is based on visual examination of a joint(s). Visual analysis can be implemented via live assessment or with a standard video recording device that enables slow motion and freeze frame capabilities.[8] Results from the current study are comparable to other previously published observational studies. Mackey et al.,[28] found K values ranging from 0.43-0.86 from video observational gait analysis in children with spastic diplegia. Children with spastic cerebral palsy have also been assessed using observational analysis with inter-observer reliability ranging between 0.59-0.79.[29] Sport specific observational analyses have also shown to successfully identify correct and incorrect freestyle biomechanics during swimming with K values ranging from 0.50-0.90.[30]

Though promising, the study had limitations. First, the sample size of observers was small and restricted to high school coaches and tennis professionals. Future research should include a variety of different coaches with different backgrounds to determine whether these factors contribute to the success of classroom training in assessing tennis serve mechanics. Second, reliability of the OTSA was only in the context of a classroom instructional training session. Future work is under way to determine if a web-based learning session yields comparable or even better reliability scores for OTSA administration.

CONCLUSION

Our findings suggest that the majority of components associated with the OTSA can be reliably taught to novice users via a classroom training session. Forward hip lean (node 5) appears difficult to assess reliably, and should be consider for removal from the final version of the tool if the suggested adaptations do not improve reliability in future studies. Our results confirm that the OTSA may be used by coaches and HCP to reliably assess tennis serve mechanics.

References

1.Johnson CD, McHugh MP. Performance demands of professional male tennis players. *Br J Sports Med*. 2006;40(8):696-699.

2.Abrams GD, Sheets AL, Andriacchi TP, Safran MR. Review of tennis serve motion analysis and the biomechanics of three serve types with implications for injury. *Sports Biomech*anics / International Society Of Biomechanics In Sports. 2011;10(4):378-390.

3.Fleisig GS, Nicholls R, Elliot BC, Escamilla RF. Kinematics used by world class tennis players to produce high-velocity serves. *Sports Biomech*anics. 2003;2(1):51-64.

4.Martin C, Bideau B, Ropars M, Delamarche P, Kulpa R. Upper limb joint kinetic analysis during tennis serve: Assessment of competitive level on efficiency and injury risks. *Scand J Med Sci Sports*. 2014;24(4):700-707.

5.Martin C, Kulpa R, Ropars M, Delamarche P, Bideau B. Identification of temporal pathomechanical factors during the tennis serve. *Med Sci Sports Exerc*. 2013;45(11):2113-2119.

6.Myers NL KW, Lamborn L, Smith BJ, English RA, Jacobs C, Uhl TL. Reliability and Validity of a Biomechanically Based Analysis Method for the Tennis Serve. *International Journal of Sport Phyisical Therapy.* 2017;12(3):437.

7.Ford KR, Myer GD, Hewett TE. Reliability of landing 3D motion analysis: implications for longitudinal analyses. *Med Sci Sports Exerc*. 2007;39(11):2021-2028.

8.Ferrarello F, Bianchi VA, Baccini M, et al. Tools for observational gait analysis in patients with stroke: a systematic review. *Phys Ther.* 2013;93(12):1673-1685.

9.Krebs DE, Edelstein JE, Fishman S. Reliability of observational kinematic gait analysis. Phys Ther. 1985;65(7):1027-1033.

10.Martin K, Hoover D, Wagoner E, et al. Development and reliability of an observational gait analysis tool for children with Down syndrome. *Pediatr Phys Ther.* 2009;21(3):261-268.

11.McClure PW, Tate AR, Kareha S, Irwin D, Zlupko E. A clinical method for identifying scapular dyskinesis, part 1: reliability. *J Athl Train*. 2009;44:160-164.

12.Ekegren CL, Miller WC, Celebrini RG, Eng JJ, Macintyre DL. Reliability and validity of observational risk screening in evaluating dynamic knee valgus. *J Orthop Sports Phys Ther*. 2009;39(9):665-674.

13.Crossley KM, Zhang WJ, Schache AG, Bryant A, Cowan SM. Performance on the single-leg squat task indicates hip abductor muscle function. *Am J Sports Med*. 2011;39(4):866-873.

14.Elliott B, Fleisig G, Nicholls R, Escamilia R. Technique effects on upper limb loading in the tennis serve. Journal of science and medicine in sport / *Sports Medicine Australia*. 2003;6(1):76-87.

15.Girard O, Micallef JP, Millet GP. Influence of restricted knee motion during the flat first serve in tennis. *J Strength Cond Res*. 2007;21(3):950-957.

16.Elliott B, Wood G. The biomechanics of the foot-up and foot-back tennis service techniques. Aust *J Sports Sci*. 1983;3(2):3-6.

17.Whiteside D, Elliott B, Lay B, Reid M. The effect of age on discrete kinematics of the elite female tennis serve. *J Appl Biomech*. 2013;29(5):573-582.

18.Reid M, Elliott B, Alderson J. Lower-limb coordination and shoulder joint mechanics in the tennis serve. *Med Sci Sports Exerc*. 2008;40(2):308-315.

19.Campbell A, Straker L, O'Sullivan P, Elliott B, Reid M. Lumbar loading in the elite adolescent tennis serve: link to low back pain. *Med Sci Sports Exerc*. 2013;45(8):1562-1568.

20.Bahamonde RE. Changes in angular momentum during the tennis serve. *J Sports Sci*. 2000;18(8):579-592.

21.Cohen J. A coefficient of agreement for nominal scales. *Educ Psych Meas*. 1960;20((1)):37-46.

22.Warrens MJ. Inequalities between multi-rater kappas. *Advances in Data Analysis and Classification*. 2010;4(4):271-286.

23.Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull*. 1971;76(5):378.

24.StatstoDO. Cohen's and Fleiss's Kappa for Ordinal Data Explained. 2016; https://www.statstodo.com/CohenFleissKappa_Pgm.php. Accessed September 24th, 2016.

25.Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159-174.

26.Goodkin R, Diller L. Reliability among physical therapists in diagnosis and treatment of gait deviations in hemiplegics. Percept Mot Skills. 1973;37(3):727-734.

27.Eastlack ME, Arvidson J, Snyder-Mackler L, Danoff JV, McGarvey CL. Interrater reliability of videotaped observational gait-analysis. Phys Ther. 1991;71(6):465-472.

28.Mackey AH, Lobb GL, Walt SE, Stott NS. Reliability and validity of the Observational Gait Scale in children with spastic diplegia. Dev Med Child Neurol. 2003;45(1):4-11.

29.Araújo P, Kirkwood R, Figueiredo E. Validity and intra-and inter-rater reliability of the Observational Gait Scale for children with spastic cerebral palsy. *Brazilian Journal of Physical Therapy*. 2009;13(3):267-273.

30.Virag B, Hibberd EE, Oyama S, Padua DA, Myers JB. Prevalence of freestyle biomechanical errors in elite competitive swimmers. *Sports Health*. 2014;6(3):218-224.

About the Authors

Natalie L. Myers PhD, ATC, PES

Department of Health & Human Performance, Texas State University, San Marcos, TX

W. Ben Kibler MD, FACSM

Shoulder Center of Kentucky, Lexington Clinic Orthopedics, Lexington, KY

Gilson J. Capilouto PhD, FASHA

Rehabilitation Sciences, University of Kentucky, Lexington, KY

Philip Westgate, PhD

Biostatistics, University of Kentucky, Lexington, KY

Tony English, PhD, PT

Rehabilitation Sciences, University of Kentucky, Lexington, KY

Tim L. Uhl PhD, ATC, PT, FNATA

Rehabilitation Sciences, University of Kentucky, Lexington, KY

Corresponding Author:

Natalie Myers

Texas State University

Email: natalie.myers@txstate.edu